**1. Student bi-weekly performance summary**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Adm. No. | Name | No. of hours present | Progress1 | Remarks |
| 1. 2222329 | Goh Rui Zhuo | 8 | A | - |
| 2. 2222291 | Toh Kien Yu | 8 | A | - |
| 3.2235189 | Wang ZiXin | 8 | A | - |

1 State whether: A=On Schedule B=Ahead Schedule for no. of days C=Behind Schedule for no. of days

**2. Bi-Weekly Scrum**

|  |  |
| --- | --- |
| Week No: 6-8 Date: 13 Nov – 17 Nov  Next Week 20 Nov – 24 Nov | |
| Member Name 1: | **Goh Rui Zhuo** |
| Last 2 weeks’ Progress | * Data Preprocessing on dataset gotten from ETL pipeline with techniques.   + Outlier detection to remove outliers in seconds data.   + Negative speed data was also removed. * Exploratory Data Analysis was also done with matplotlib.   + Sensor, Driver, and safety labels was analysed.   + Analyse on merged data. * Creation of 2 SQL query that analyse on key important aspect.   + First query to analyse on car age with respect to the percentage of safety labels to provide a comprehensive analysis to see if car age affects the safety labels. |
| This 2 weeks deliverables | * Improve on Data Analysis Imputation   + Try and find out other imputation methods that is a more algorithmic than using the mean of each column by booking ID. * Improve on Exploratory Data Analysis   + Switching to a more interactive graph library with the use of plotly.   + Convert to the use of subplots to provide a cleaner visualisation.   + Adding the plotly dash to provide more functionality * Tableau Creation   + Start on dashboard creation that satisfy the user stories.   + User Interface must be vsual appealing * Checking of all work   + Double check on work before submitting |
| Obstacles | Struggle with creating two queries but ended up with query that resulted in key insights and provide valuable idea |
| Member Name 2: | **Toh Kien Yu** |
| Last 2 weeks’ Progress | * Data cleaning   + Imputed missing values in the dataset   + Dropped extreme outliers and values from the dataset such as extreme values of seconds. * Exploratory data analysis on dataset   + Merging of datasets for analysis between tables   + Created visualizations on sensors, drivers and safety label dataset * Added markdowns and analysis on charts * Created SQL queries for meaningful insights   + Created query 2 to analyse on the danger percentage of various trip durations and how it impacts on safety labels. * Refine Product Backlog |
| These 2 weeks deliverables | * Building of dashboard in Tableau   + Generate Tableau charts and choose the insightful chart corresponding to the user stories.   + Build the dashboard for manager for analysis   + Make sure dashboard are relevant, insightful, and user-friendly to corresponding users. * Dashboard interactivity   + Add filter options for users to display specific information they desire. * Project Report   + Add sub-headers   + Add Report Content * Presentation Slides   + Create a deck of slides and document down the procecss * Rechecking of all the deliverables   + Reviewing all deliverables to confirm that all have been finished entirely. |
| Obstacles | * Had difficulty crafting out the formula/logic behind the calculated field in Tableau. |
| Member Name 3: | **Wang ZiXin** |
| Last 2 weeks’ Progress | * Creation of ERD diagram   + Check the data type and relation type. * Creation of user story   + Acceptance criteria and conformation * Data preprocessing   + Remove the unwanted column from dataset * EDA   + Creating of visual charts (histograms, BarCharts, scatterplot) * Creating SQL queries   + Create a query to find relationship between age group of drivers and rating |
| This 2 weeks’ deliverable | * Tableau Creation   + Create dashboard and sheets to fit the confirmation * Report   + Add content into report * Creation of PowerPoint slides * Double check   + Double check everything after submission * Improve 1 sql query to get better insite |
| Obstacles | Take long time to create nice sheets |